Working Paper 1.2

Overview of Assessment Methods and Model Selection Criteria for Nineteen Groundfish Stocks in the Northeast US

A Working Paper in Support of Term of Reference 1

GARM 2008 Models Meeting
Woods Hole, MA
25-29 February 2008

Background

This report provides a summary of the assessment approaches that have been used to date for the 19 GARM stocks. Each section gives a brief overview of the methods that have been used to date, an evaluation of their strengths and weaknesses, and the prospects for implementing alternative approaches. This report complements the data and analyses summarized in Working Paper 1.1. To reduce the size of this report,   Tables and Figures referenced in this report are in Working Paper 1.1.

Model selection requires simultaneous consideration of multiple objectives. Alternative models should be technically superior in terms of modern statistical theory, flexibility to handle heterogeneous data and processes, stability of results, and expanded scope of inference. Alternative models must also allow for the provision of scientific advice with respect to reference points. One aspect of particular importance in the Northeast US is stock rebuilding. The model must include a forecasting component or be linked with forecasting program. This is necessary for the evaluation of alternative management strategies. Such evaluations are also required as components of the economic impacts of alternative harvest levels. Finally, coordination of assessment approaches with states and Canada is important for jointly-managed species.

Recommendations for model selection will reflect a compromise of these multiple objectives and constraints. These summaries are designed to facilitate discussions on final recommendations for model selection (TOR 5) by providing necessary background on approaches previously used and new approaches.   It is anticipated that the results of Working Papers for TOR 2 and 4 will be particularly informative for model selection.

A. Georges Bank Cod by Loretta O'Brien

1. Description and history of current approach

The ADAPT calibration method (Parrack 1986, Gavaris 1988, Conser and Powers 1990) has been used in VPA to assess Georges Bank cod since 1986 (Serchuk and Wigley 1986). In the most recent assessment (O'Brien et al. 2006), landings at age consisted of combined USA and Canadian landings from 1978-2004 for ages 1-10+. The indices of abundance used to calibrate the VPA included the NEFSC 1978-2005 spring indices for ages 1-8 and autumn indices for ages 0-6, and the DFO 1986-2005 spring indices for ages 1-8. Fully recruited F occurs at age 4, and F on the oldest age is an average of ages 4-8.

2. Key Strengths and Weaknesses

The available data for Georges Bank cod i.e. commercial landings that are relatively well sampled, and three research surveys, are very applicable to an age-structured model such as VPA. Discard at age data, when available, can easily be included in the model. Total estimates of recreational catch are available. Their use in VPA requires that they be disaggregated by age using and age-length key from other sources. Until recently direct biological sampling of recreational cod fisheries has been infrequent. Models that can utilize total recreational catch can be useful. Prior to 1994, when CPUE was used in the calibration, and prior to significant management measures (e.g.year round closed areas, TACs, change in sampling regimes) the VPA performed well. Since 1994, consistent retrospective patterns have persisted in estimates of F and SSB. This may be in part due to changes in selectivity, however, shifts in selectivity can not be parameterized directly in the VPA model.

3. Feasibility of changing assessment models

Preliminary analyses and comparison of SS2 and VPA (see TOR 2 WP2.7 Jacobson et al.) suggests that a forward projecting model, such as SS2 or ASAP, may be feasible for this stock. Sufficient data is available for these models and the software is readily available. SS2 is not currently programmed, however, to produce BRP as estimated at the NEFSC. GB cod is a transboundary stock, and both the NEFSC and DFO currently assess the stock using VPA. Any change in models by the NEFSC would introduce an inconsistency between the stock assessments.

References

Conser, R.a.J.E.P. 1989. Extensions of the ADAPT VPA tuning method designed to facilitate assessment work on tuna and swordfish stocks. ICCAT Working Doc. SCRS/89/43 Revised 15 p.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. CAFSAC Res. Doc 88/29 12 p.

O'Brien, L., N.Shepherd, and L. Col 2006. Assessment of the Georges Bank Atlantic cod stock for 2005. Northeast Fisheries Science Center Ref. Doc. 06-10 148 p.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. Conserv. Atlantic Tunas. Coll. Vol. Sci. Pap. **24**: 209-221.

Serchuk, F. M and S. E. Wigley. Assessment and status of the George Bank and Gulf of Maine Atlantic cod stocks - 1986. Woods Hole Lab. Ref. Doc. No. 86-12 . 86.

B. Georges Bank Haddock—Liz Brooks

1. Current Approach

Georges Bank haddock is currently assessed using a VPA that incorporates catch at age (ages 1-9^{+}) and indices from 1963 to 2006. A total of 30 indices were used in the previous VPA (Brodziak et al. 2006): NEFSC spring survey for ages 1-8 (Yankee 36 net), NEFSC spring survey for ages 1-8 (Yankee 41 net, years 1973-1981), NEFSC fall survey for ages 0-5 (lagged 1 year), and Canada spring survey for ages 1-8. No retrospective pattern has been observed for this model formulation in the past decade, which likely explains why alternative models have not been applied to this stock. Reference points were derived by a post-hoc ("non-parametric") analysis of VPA estimates of spawning biomass and recruits. This approach identified a "threshold" level of SSB that separated two apparent "average" recruitment levels (Brodziak et al. 2001).

2. Key strengths and weaknesses

The base case VPA has only been applied with a starting year of 1963 because that is when the surveys begin. However, catch at age is available as far back as the 1930s (Clark et al. 1982). An alternative model that did not require complete overlap between catch at age and indices would allow incorporation of more historic landings information. Also, the lack of an internal link in the VPA between spawning biomass and recruitment may predispose the estimated trajectories to appear unrelated even if they truly were. In other words, there is a logical inconsistency in analyzing VPA estimates of SSB and recruitment (which were derived assuming no relationship) to determine if there is a stock-recruit relationship. An alternative model that had the option to estimate a stock recruit function would provide consistency between the model and estimated reference points.

3. Feasibility of changing assessment model

All stock assessment and management to date has been based on VPA results. The data available for this stock would be appropriate to analyze in any of the forward projecting statistical catch at age models, such as ASAP, STATCAM, or SS2. However, STATCAM at present only estimates an average level of recruitment with annual deviations, and as such would only be able to estimate reference points based on a YPR proxy rather than MSY or $SPR_{\%}$ reference points derived from a stock-recruit function.

Literature cited

Brodziak, J.K.T., W.J. Overholtz, and P.J. Rago. 2001. Does spawning stock affect recruitment of New England groundfish? Can. J. Fish. Aquat. Sci. 58:306-318.

Brodziak, J.K.T., M. Traver, L. Col, and S. Sutherland. 2006. Stock assessment of Georges Bank Haddock, 1931-2004. NEFSC CRD 06-11.

Clark, S.H., W.J. Overholtz, and R.C. Hennemuth. 1982. Review and assessment of the Georges Bank and Gulf of Maine haddock fishery. J. Northw. Atl. Fish. Sci. 3:1-27.

C. Georges Bank Yellowtail Flounder by C.M. Legault

1. Current Approach

Georges Bank yellowtail flounder is currently assessed using a VPA that splits all the survey abundance at age series at 1995, called the "Major Change" model. This approach reduces the strong retrospective pattern observed in the "Base Case" VPA, that does not split the survey time series. However, the Major Change model results in three to five fold increases in estimated survey catchabilities. These changes in catchability cannot be explained by changes in the survey design or operations, but are instead considered to be aliasing an unknown mechanism which produces the retrospective pattern in the Base Case VPA. Statistical catch-at-age models have been applied in the past and demonstrate similar retrospective patterns when the survey time series are kept whole. Surplus production models have also been applied in the past and consistently produce strongly increasing trends in biomass in recent years that do not agree with survey trends, and so have not been used for management advice.

2. Key strengths and weaknesses

A benchmark assessment for Georges Bank yellowtail flounder was conducted in 2005 with many outside experts in an attempt to resolve the issue with the retrospective pattern observed in the Base Case VPA. A wide range of modeling approaches were examined, but no resolution could be reached regarding the basic problem of insufficient catch to cause the recent decline in the surveys and the relative lack of old fish. The Major Change VPA was proposed as an alternative to the Base Case VPA, but both models were used for management advice, because the large increase in estimated survey catchabilities could not be explained. Since then, the Transboundary Resources Assessment Committee (TRAC) has concluded that the retrospective pattern in the Base Case VPA is too strong to allow its use for management advice and so uses only the Major Change VPA for management advice. The inability to explain why the Base Case VPA (or any other model with whole survey time series) has a retrospective pattern or else obviously incorrect population dynamics remains a major weakness in this assessment.

3. Feasibility of changing assessment model

This assessment has been and is currently being conducted using multiple software packages (VPA, ASPIC, ASAP, SS2). Data do not necessarily limit any of the modeling approaches for this stock. However, full utilization of length data, especially if separated into males and females, or fleet-based catch data would require some time.

D. Southern New England-Mid Atlantic Yellowtail Flounder by C.M. Legault and S.X. Cadrin

1. Current Approach

Southern New England-Mid Atlantic yellowtail flounder is currently assessed using a VPA, with a surplus production model providing confirmatory evidence of stock status. Both models indicate a severely depleted stock in recent years. The VPA estimates poor recruitment every year since 1991 and a high fishing mortality rate, despite substantial reductions in catch in the last decade. Low landings in recent years caused difficulties in collecting port samples. An Industry-Based Survey (IBS) was conducted from 2003 to 2005 and has provided much of the length and age information for the recent catch.

2. Key Strengths and Weaknesses

The close collaboration with the fishing industry through the IBS has lead to significant improvement in characterization of the catch at age in recent years. The ASPIC model has problems due to the extended period of low abundance and recruitment in recent years because it expects high productivity at low abundance. The VPA has been found to be quite sensitive to the age of the plus group and the assumption made when deriving F on the oldest true age. Accurate estimates of total catch and characterization of catch at age will continue to be a challenge, because the IBS was discontinued, landings remain low and difficult to sample, and a substantial portion of the catch is discarded.

3. Feasibility of changing assessment model

Data do not limit the application of any modeling approach for this stock. Age data are available for all components of the catch at age (landings and discards). Data do not necessarily limit any of the modeling approaches for this stock. However, full utilization of length data, especially if separated into males and females, or fleet-based catch data would require some time.

E. Cape Cod-Gulf of Maine Yellowtail Flounder by C.M. Legault, S.X. Cadrin and J. King

1. Current Approach

Cape Cod-Gulf of Maine yellowtail flounder is currently assessed using a VPA. The results exhibit a moderate retrospective pattern. Similar to the Georges Bank yellowtail flounder, there is a relative lack of old fish in both the catch and surveys that result in high F estimates from the VPA, even though the catches appear modest. The most recent assessment, GARM 2005, showed a decline in recent recruitment and spawning stock biomass as well as an increasing trend in F.

2. Key Strengths and Weaknesses

The Massachusetts Department of Marine Fisheries conducts spring and fall surveys which catch large numbers of Cape Cod-Gulf of Maine yellowtail flounder in addition to the NEFSC spring and fall surveys. A new survey time series is available from the Maine-New Hampshire region which may prove useful for this assessment. Port sampling has improved in recent years, which has decreased the uncertainty in landings at age estimates. The VPA is quite sensitive to the age of the plus group and the assumption made when deriving F on the oldest true age, similar to the Southern New England-Mid Atlantic yellowtail flounder assessment.

3. Feasibility of changing assessment model

Catch at age information is available for all components of the fishery (landings and discards), although the discard estimates need to be updated. Data do not necessarily limit any of the modeling approaches for this stock. However, full utilization of length data, especially if separated into males and females, or fleet-based catch data would require some time.

F. Gulf of Maine Cod

1. Description and history of current approach.

The Gulf of Maine stock of Atlantic cod has been assessed since 1986 (Serchuk et. al. 1978), but the first age-disaggregated assessment using Virtual Population Analysis was not accomplished until the early 1990s when a sufficient time series of catch-at-age data became available. The most recent peer review occurred at SAW 31 in 2001 (Mayo et al. 2002). Since then updated assessments were reviewed at GARMI (2002) and GARMII (2005) Mayo and Col (2006). The VPA is calibrated using the Adapt method (Parrack 1986, Gavaris 1988, Conser and Powers 1990).

2. Key strengths and weaknesses

VPA has been successful in capturing the large-scale stock dynamics of the Gulf of Maine cod stock (i.e., trends in fully recruited fishing mortality and stock size) throughout most of the period since 1982. In recent years, estimates of F have been declining and with Fs projected to decline to Fmsy or less, the ability of VPA to reflect stock dynamics may be diminished. The age composition for this stock include true ages 1-6 and a 7+ group. The first fully recruited age is age 4. For logistical purposes, F on the oldest true age (age 6) is based on only two ages (ages 4 and 5). Estimates of terminal F, therefore, can be variable from year to year and highly dependent on a single value, as was the case in the 2005 GARMII assessment. In this case the estimate of the 2004 F on age 4 was 0.78 and the estimate on age 5 was 0.38, resulting in an unweighted average fully recruited F of 0.58. This value reversed a trend in declining F evident for the previous 6 years. While this may have been the result of a very weak year class (2000) entering the fishery (estimates of stock size were not as severely affected), it is uncertain if this event will recur in the future. Although there is no consistent retrospective **pattern** for this stock, F and stock size are generally over- or under estimated in the terminal year.

3. **Feasibility of changing assessment model**

Alternative forward projecting approaches for assessing this stock can use the existing age data in the same manner as VPA These provide direct estimates of biological reference points and can provide the same input data to NOAA/NFT projection software. They have been validate and are practical to implement.

G. Witch Flounder

1. Description and History of current approach:

Witch flounder is assessed as a unit stock. This species is slow-growing, late-maturing, and long-lived. Analytic assessments (VPA) have been conducted for witch flounder since 1994 using catch (landings and discards) and NEFSC spring and autumn survey indices to estimate stock sizes for ages 3 to 10, average fishing mortality (ages 8-9 unweighted) and spawning stock biomass.

In 2003, age-structured forward-projection model (statistical catch at age - SCAA) was fit to fishery and survey data during 1937 to 2002. This model provided an alternative long-term perspective on resource dynamics in comparison to VPA-based analyses that were limited to 1982 to 2002. There was general agreement between VPA and SCAA results during 1982 to 2002. Spawning stock biomass estimates were very similar during 1989-1999; however, the VPA indicated smaller decreases in spawning stock biomass during 1982-1988 and a greater increase during 2000-20002. Fishing mortality estimates were also similar; both VPA and SCAA estimates increased until the mid-1990s and then declined. Recruitment estimates also exhibited similar patterns, although the VPA indicated larger increases in recruitment during the late-1990s. Despite differences in model configuration and estimation approaches, the SCAA generally confirmed point estimates and trends in the VPA results. See Wigley et al. (2003) for further details on this analysis.

Witch flounder was last assessed in 2005 using a VPA model. See the witch flounder section in Mayo and Terceiro (2005) for further details on the VPA analyses. The retrospective analyses indicates that average F was underestimated in the late 1990s and early 2000s, spawning stock biomass was consistently overestimated, and a pattern of relatively consistent estimates of the number of age 3 recruits, with the notable exception of the 1992, 1993 and 1996 year classes, which were considerably overestimated. Bootstrap results suggest that the estimates of F and spawning stock biomass are relatively precise with CVs of 30% and 14%, respectively.

Based on yield and spawning stock biomass per recruit analyses and the arithmetic mean of the VPA age 3 recruitment, the biological reference points are:

SSBmsy = 25,248 mt
Fmsy = F40% = 0.23
MSY = 4,375 mt.

2. Key strengths and weaknesses:

Strengths: NEFSC survey trends are corroborated by Mass. inshore survey and Atlantic States Marine Fisheries Commission summer shrimp survey.

Weaknesses include: 1) low frequency of samples across some market categories and quarters results in imprecise mean weights at age and estimates of numbers at age in the early 1990s (sampling has since improved); 2) lack of data to support direct estimates of discards at age requires use of a surrogate survey-based methods; 3) retrospective patterns suggest that estimates of SSB may be overestimated (e.g. future assessments may have lower estimates of SSB); 4) the research bottom trawl survey catches very few witch flounder; in most years, the stratified mean number per tow of witch flounder is less than 5 fish. Abundance of witch flounder in the late 1980s and early 1990's may have gone below levels that provide reliable estimates of trends in abundance and biomass.

3. Feasibility of changing assessment model:

Age data: commercial and survey age data will be available through 2007.
Validated software: VPA software is available in the Toolbox.
Linkage with biological reference points: Linkage between VPA and yield and spawning stock biomass per recruit analyses are available.
Linkage with projections: linkages between VPA and projection software are available.

References:

Mayo, R.K and Terceiro, M. editors, 2005. Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole MA, 15-19 August 2005. *U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc.* 05-13, 499 p.

Wigley, S.E, Brodziak, J.K.T. and Col, L. 2003. Assessment of the Gulf of maine and Georges Bank withc flonder stock for 2003. Northeast Fisheries Science Center Reference Document 03-14, 186 p.

H. Gulf of Maine/Georges Bank American Plaice--Loretta O'Brien

1. Description and history of current approach

The ADAPT calibration method (Parrack 1986, Gavaris 1988, Conser and Powers 1990) has been used in VPA to assess Gulf of Maine - Georges Bank American plaice since 1982 (O'Brien et al). In the most recent assessment (O'Brien et al. 2005), catch at age consisted of combined landings and discards from 1980-2004 for ages 1-9+. The indices of abundance used to calibrate the VPA included the 1980-2005 NEFSC spring indices for ages 1-8 and autumn indices for ages 1-6, and the 1982-2005 MADMF spring and autumn indices for ages 1-5. Fully recruited F occurs at age 5, and F on the oldest age is an average of ages 5-8.

2. Key Strengths and Weaknesses

The available data for the Gulf of Maine-Georges Bank American plaice i.e. combined commercial landings that are relatively well sampled, discard catch at age for two fleets, and four research surveys, are very applicable to an age-structured model such as the VPA. The VPA for this stock has performed well, with relatively little retrospective pattern, but only with the inclusion of discards at age. The assessment was conducted initially with both areas combined primarily due to sampling, however, analyses subsequently showed that the growth of American plaice in the Gulf of Maine is significantly slower than that of the fish on Georges Bank (Esteves and Burnett 1993) There is also likely dimorphic growth that is not easily accounted for given the current sampling protocol.

3. Feasibility of changing assessment models

Sufficient data is available for a forward projecting model such as SS2 or ASAP and the software is readily available. These models would need to produce/provide data for estimation of projections and BRPs. An evaluation of the ADAPT VPA vs. other models would need to be conducted to assess the feasibility of changing models. A model that could account for the differences in growth between GM and GB would be a potential improvement. A more logical first step may be to re-estimate the landings at age by sub-region

Reference List

Conser, R. and J. Powers. 1990. Extensions of the ADAPT VPA tuning method designed to facilitate assessment work on tuna and swordfish stocks. ICCAT Working Doc. SCRS/89/43 Revised 15 p.

Esteves, C. and J. Burnett 1993. A comparison of growth rates for American plaice, *Hippoglossoides platessoides*, in the Gulf of Maine-Georges Bank region derived from two different data sources. NEFSC Ref. Doc. 93-09. 8.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. CAFSAC Res. Doc 88/29 12 p.

O'Brien, L., J. Burnett, and L. Col. 2005. H. Gulf of Maine - Georges Bank American Plaice *in*: Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 15-19, 2005. NEFSC Ref. Doc. 05-13. 508p.

O'Brien, L., R.K. Mayo, N.Buxton, and M.Lambert. Assessment of American Plaice in the Gulf of Maine-Georges Bank Region - 1992. Appdx CRD-92-07. Res. Doc SAW 14/2 . 82.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. Conserv. Atlantic Tunas. Coll. Vol. Sci. Pap. **24**: 209-221.

I. Gulf of Maine Winter Flounder –Paul Nitschke

1. Description and history of current approach.
VPA modal was first used in SARC 36 and GARM I in 2001 and updated in GARM II. In GARM III the VPA was updated with re-estimated commercial landings, large mesh trawl and gillnet discards, and recreational catch through 2004. The use of the SCALE and AIM model with data through 2006 were also investigated.

2. Key strengths and weakness.
The scale model does not require production aging and has some increased flexibility in model configuration compared to the VPA. However, all of the model examined suffer from the apparent lack of the relationship between the trends in catch and the abundance indices. The lack of dynamics in the size and age distributions over time will be difficult to explain with the large decrease in catch in regardless of the model framework.

3. Feasibility of changing assessment model.
Investigating the use of SS2 or ASAP as an alternative age based forward projection model can be considered. However the use of alternative model will likely suffer from the same issues seen in the SCALE and VPA models without a large change in assumptions. The present SCALE and VPA model configurations are presently both producing unreliable results.

J. Southern New England Winter Flounder—Mark Terceiro

1. Description and history of current approach.

This stock is assessed with VPA ADAPT as formulated in the NOAA Fisheries Toolbox. Commercial and recreational fishery landings and discards estimates at age, and the total fishery catch at age used as input. The primary gear in the fishery is the otter trawl which accounts for an average of 98% of landings since 1989. Recreational landings are important, typically constituting about 10% of commercial landings in recent years

Length samples of winter flounder are available from both the commercial and recreational landings. In the commercial fishery, annual sampling intensity varied from 23 to 251 mt landed per 100 lengths measured during 1981-2006.

Prior to 1994, NEFSC trawl survey length frequencies and commercial trawl fishery mesh selection data were used to estimate the magnitude and characterize the length frequency of the commercial fishery discard. For 1994-2006, NEFSC Fishery Observer trawl and scallop fishery discards to landings ratio estimates were applied to corresponding commercial fishery landings to estimate discards in weight. A discard mortality rate of 50% was applied to commercial fishery live discards; recreational live discard mortality rate is assumed to be 15%

The following NEFSC and state abundance indices at age are used in VPA calibration: NEFSC spring, fall, and winter bottom trawl, the Massachusetts spring trawl survey, Rhode Island fall seine survey, Rhode Island spring trawl, Connecticut spring trawl, New York trawl, Massachusetts summer seine, Delaware juvenile trawl, New Jersey Ocean trawl, and New Jersey River trawl surveys .

2. Key strengths and weakness.

Despite the large number of fishery independent surveys and availability of commercial landings and discard information the SNE/MA winter flounder VPA exhibits a severe retrospective pattern. Underestimation of F and overestimation of SSB occurs and is most likely caused by underestimation of the total catch. The overestimation of SSB was most severe for the 1997 and 1998 terminal years. The retrospective estimation of age-1 recruits indicated a tendency for overestimation during 1993-2001.

No alternative forward projection models have been explored for this stock.

3. Feasibility of changing assessment model.

The wealth of landings, discard and survey data could certainly support development of a forward projection model. Revisions of the assessment information for this purpose would be difficult in view of the demands for assessments of higher priority stocks within the GARM. Experiences with Gulf of Maine winter flounder suggest that the retrospective pattern could pose difficulties for both SS2 and ASAP.

K. Georges Bank Winter Flounder—Lisa Hendrickson

1.0 Assessment History and Proposed Modeling Approach for GARM 2008

The assessment history of the Georges Bank winter flounder stock, model input data, and model performance are summarized in Table K1. The current modeling approach involves an ASPIC production model because the results of the VPA runs presented at SARC 34 were considered unreliable, primarily due to inadequate fishery length and age samples during the terminal years of the assessment period. Weaknesses of using an ASPIC model are the loss of important information regarding size and age composition, the inherent re-estimation of reference points with each model run, and the fact that model estimates of relative total biomass (B_t/B_{MSY}) and fishing mortality rates (F_t/F_{MSY}) are more precisely estimated than the absolute values. The latter fact is problematic with respect to making stock status determinations which are based on absolute reference points. Strengths of the use of an ASPIC model include a reasonable fit with no strong retrospective pattern in the fishing mortality or biomass estimates.

The sampling intensity of fishery length and age composition data has improved greatly since 2001, when the most recent VPA was conducted. Therefore, both VPA and ASAP models, for 1982-2007, are proposed for the 2008 GARM.

Table K1. Assessment history of the Georges Bank winter flounder stock.

Assessment Venue	Model Type	Input Data	Model Performance
SAW 28 (Fall 1998)	VPA	1982-1997 NEFSC spring and fall surveys (offshore strata 13-22); door conversion factors (weight = 1.39, numbers = 1.46 for 1982-1985)	1990-1997 retrospective patterns (slight overestimation of F and slight underestimation of SSB during 1991-1992 then overestimated after 1993; tendency to underestimate recruitment in terminal year)
		1987-1997 CA Feb. surveys (5Z1-5Z8); use NEFSC spring survey A/L key M = 0.2 1982-1998 U.S. LAA and WAA (ages 1-7+),SA 522,525,561,562; CA landings (SA 551,552) assumed to have same age comp. as U.S. landings 1982-1998 maturity ogive from NEFSC spring surveys	
SAW 34 (Fall 2001)	VPA	1987-2000 NEFSC survey data same as SAW 28 (NEFSC spring, ages 1-7; NEFSC fall ages 3-6 (lagged 1 yr and age)	<u>Not adopted</u> due to high CV range for stock size at ages 2-6 (28-60%) and 1993-2000 retrospective patterns (F underestimated in terminal year but no patterns in SSB or recruitment); inconsistent patterns in mean weights at age and Fs at age (post-1993)
		1987-2000 CA survey data same as SAW	

		28, but included strata 5Z1-5Z4 (ages 4-7) M, maturity ogive, WAA, LAA same as SAW 28 but LAA and WAA for 1-10+	
	ASPIC Production Model	Relative biomass indices for: NEFSC fall surveys (1964-2000, lagged forward 1 yr and age) 1968-2001 NEFSC spring surveys Total landings (US, CA, and USSR)	(Accepted model) Reasonable model fit; trends in avg. B and F similar to VPA trends; no retrospective patterns in F or SSB (1995-2000)
	STATCAM (forward-projecting age-structured model)		Exploratory only
GARM (Fall 2002)	ASPIC Production Model	Update of SAW 34 formulation of ASPIC model using data through 2001	Reasonable model fit
GARM (Fall 2005)	ASPIC Production Model	Update of SAW 34 formulation of ASPIC model using data through 2004	Reasonable model fit with no strong retrospective patterns in F or avg. B

L. Georges Bank/Gulf of Maine White Hake –Katherine Sosebee

1.0 Current Approach

The assessment for this stock has evolved over time from index-based in the early 1990s, to Collie-Sissenwine in 1994, finally to VPA in 1998. However, the addition of years to the VPA model created a significant retrospective pattern in the assessment in 2001. The assessment then became a surplus production model which was itself unstable and rejected in 2002. The AIM method is currently used to assess the status of the stock relative to biological reference points. Landings and discards of fish greater than 60 cm, are used in the model as well as autumn survey indices of biomass.

2.0 Key Strengths and Weaknesses

The strength of the AIM method is in its simplicity. The use of catch and survey data of animals greater than 60 cm ensures that only information about white hake is included in the model. One weakness is that in using information about the large size classes, only part of the population is represented in the current AIM model. Another weakness is that other sources of data (other surveys, age and length compositions) are currently not used in the modeling efforts.

3.0 Feasibility of Changing Assessment Model

The current data limitations are the lack of age information in the most recent years. If a VPA were to be used these data would be required and updating would require some time. The use of a model such as SS2, ASAP, or SCALE would not require these data and other sources of information could be used.

M. Georges Bank/Gulf of Maine Pollock—Ralph Mayo

Pollock, *Pollachius virens* (L.) have traditionally been assessed as a unit stock from the Scotian Shelf (NAFO Divisions 4VWX) to Georges Bank, the Gulf of Maine and portions of the Mid-Atlantic region (Subareas 5 and 6). This stock was last assessed over its range *via* VPA at SAW 16 in 1993 (Mayo and Figuerido 1993, NEFSC 1993a, 1993b). At that time, spawning stock biomass had been declining since the mid-1980s, and was expected to reach its long-term average (144,000 mt). Fishing mortality was estimated to be 0.72 in 1992, above F20% (0.65) and well above Fmed (0.47). The stock was then considered to be fully exploited and at a medium biomass level.

The state of this stock was first evaluated *via* index assessment in 2000 (Mayo 2001). At that time, it was noted that biomass indices for the Gulf of Maine-Georges Bank portion of the stock, derived from NEFSC autumn bottom trawl surveys, had increased during the mid-1970s, declined sharply during the 1980s, but have been generally increasing since the mid-1990s. Indices derived from Canadian bottom trawl surveys, conducted on the Scotian Shelf, increased during the 1980s, but declined sharply during the early 1990s.

In 2002, index-based biological reference points were developed for a portion of the pollock stock primarily under US management jurisdiction (Subareas 5 and 6), including a portion of eastern Georges Bank (Subdivision 5Zc) that is under Canadian management jurisdiction (NEFSC 2002). The most recent assessment of the resource inhabiting the area comprising this management unit was conducted in October, 2005 at the second Groundfish Assessment Update Meeting (GARM II). At that time it was determined that the index of current biomass was greater than ½ of the Bmsy proxy reference point and that the index of current F was below the Fmsy proxy reference point (Mayo and Col 2002).

Strengths and Weaknesses

The current assessment is based on An Index Model (AIM) which incorporates age-aggregated information on exploitation (commercial landings) and resource biomass (Autumn NEFSC biomass index). An age-based assessment may be possible for this region if the age structure of the commercial landings can be compiled. This must include the non-USA components (Canadian and DWF) that were substantial since the early 1960s.

References

Mayo, R.K. and B.F. Figuerido. 1993. Assessment of Pollock, *Pollachius virens* (L.), in Divisions 4VWX and Subareas 5 and 6, 1993. NMFS, Northeast Fisheries Science Center Reference Document 93-13, 108 p.

Mayo, R.K and L. Col 2002. Scotian Shelf-Georges Bank-Gulf of Maine Pollock. In: Assessment of 20 Groundfish Stocks through 2001. A Report of the Groundfish Assessment Review Meeting (GARM), Northeast Fisheries Science Center Reference Document 02-16.

NEFSC 1993a. Report of the 16th Northeast Regional Stock Assessment Workshop (16th SAW). Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NMFS, Northeast Fisheries Science Center Reference Document 93-18, 118 p.

NEFSC 1993b. Report of the 16th Northeast Regional Stock Assessment Workshop (16th SAW). The Plenary. NMFS, Northeast Fisheries Science Center Reference Document 93-19, 57p.

NEFSC 2002. Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish, NMFS, Northeast Fisheries Science Center Reference Document 02-04, 254 p.

N. Gulf of Maine/ Georges Bank Acadian Redfish—Tim Miller

1. Current Approach

In the last assessment two versions of the statistical catch at age model (RED and STATCAM) were explored. Mayo *et al*. (2002) and STATCAM (2005) provide full descriptions of RED and STATCAM models, respectively.
Stock assessment results were ultimately based on the RED model.

2. Key strengths and weaknesses

We estimated assessment parameters using RED and STATCAM models with the same data as Mayo et al. 2005 and updated landings and survey data through 2006

Base RED and STATCAM models that use the same likelihood weighting factors as Mayo et al. (2005) were found to provide biologically unrealistic trends in annual biomass estimates when landings between 1913 and 1933 were included Because the weightings gave unreasonable population trends when small annual landings prior to 1934 were included, there may be reason to question that the weightings are appropriate for the shorter time series of landings. Both RED and STATCAM models are sensitive to different weightings that yield biologically plausible trends in biomass and recruitment and reasonable residual patterns for the various data components . The models are also sensitive to the minimization phase of the (log) fishing mortality parameters for (log) survey selectivity parameters .

The Alternative 1 models for RED and STATCAM give more reasonable trends in spawning biomass and (log) recruitment estimates, but for RED, the biomass estimate in 1934 is less than half the corresponding estimate for the base RED model using the shorter landings time series FOR STATCAM Alternative 1 the 1934 spawning biomass estimate is similar to the corresponding estimate for the base model and shorter landings time series, but the current (2006) biomass estimate is about 3 times the corresponding estimate with the shorter landings series.

STATCAM Alternatives 2 and 3 show better trends in spawning biomass estimates than the base models when fitting the longer times series of landings and have similar biomass estimates in the earliest years and the current biomass estimates for these models are similar to the corresponding estimates for the base models using the shorter times series of landings Both of these models do not show any strong patterns in the (log) residuals from the landings, CPUE and survey biomass per tow estimates and exhibit as good or better fits to the proportions at age in the landings and surveys than the base RED model. Estimates of spawning biomass for the 1930s using Alternatives 2 and 3 are more in line with the corresponding estimates from Mayo et al (2005).

3. Feasibility of changing assessment model

The feasibility of alternative forward projection models was demonstrated in Working Paper 1.1. Both RED and STATCAM have been used to estimate assessment model parameters. Fishing mortality rates on redfish appear to be low. This complicates the ability to detect change in any model.

Further development of the FSCTPD model would be necessary to allow estimation of total biomass and recruitment earlier than 1969 when the landings age measurements began.

O. Ocean Pout—Susan Wigley

1. Description and History of current approach:

Ocean pout is assessed as a unit stock. There is little market demand for this species; in recent years, discards exceed landings. Trends in abundance have been declining since 1980.

An index assessment has been conducted for this species since 1990 using NEFSC spring survey and landings. The status of stock is evaluated using a three year moving average biomass (kg/tow) and an exploitation ratio (landings / three year average spring survey biomass).

In 2002, an Index Asssement Model (AIM) was explored using the replacement ratio analysis. Given the weak relationship between the replacement ratio and relative F, as indicated by the circular shape of the ellipse (Figure O.1), the input data for this stock may be imprecise. It was concluded that these analyses were not informative upon which to based recommendations for Bmsy, Fmsy and MSY. Ocean pout stock status determination continues to use survey biomass and relative exploitation.

2. Key strengths and weaknesses:

Strengths: NEFSC survey trends are corroborated by Mass Inshore survey.

Weaknesses: Limited biological data (lengths and ages) from commercial fishery; biological reference points continue to be based on survey biomass and relative exploitation.

3. Feasibility of changing assessment model:

Age data: No production ageing of this species (survey or commercial)
Validated software: N/A
Linkage with biological reference points: Bmsy proxy are based on NEFSC spring survey kg/tow (1980 – 1991 median = 4.9 kg/tow); MSY chosen to be 1,500 mt; Fmsy proxy = 0.31.

P. Gulf of Maine/Georges Bank Windowpane Flounder—Lisa Hendrickson

Assessment History and Proposed Modeling Approach for GARM 2008

The Gulf of Maine-Georges Bank (GOM-GB) windowpane flounder stock has never been assessed through the Stock Assessment Review Committee (SARC) process. The stock was initially assessed in 2000 at the request of the NEFMC Multispecies Monitoring Committee and has been assessed at each GARM thereafter. With the exception of 1999, survey age composition data are unavailable as is the age composition of the landings. However, fishery and survey length data are available.

The current assessment approach is index-based. Medium term stochastic projections have been previously generated using ASPIC software and AIM has been used to derive biological reference points for the stock. Weaknesses of using the current assessment approach are the inability to incorporate the existing size composition information from the surveys and fishery. Therefore, the Collie-Sissenwine Analysis (CSA) model, which estimates the abundance and fishing and total mortality rates of recruit and post-recruit size classes, will be explored at the 2008 GARM in addition to an index-based assessment approach.

Q. Southern New England – Mid-Atlantic Windowpane Flounder-- Lisa Hendrickson

Assessment History and Proposed Modeling Approach for GARM 2008

The Southern New England-Mid-Atlantic Bight (SNE-MAB) windowpane flounder stock has never been assessed through the Stock Assessment Review Committee (SARC) process. The stock was initially assessed in 2000 at the request of the NEFMC Multispecies Monitoring Committee and has been assessed at each GARM thereafter. With the exception of 1999, survey age composition data are unavailable as is the age composition of the landings. However, fishery and survey length data are available.

The current assessment approach is index-based. Medium term stochastic projections have been previously generated using ASPIC software and AIM has been used to derive biological reference points for the stock. Weaknesses of using the current assessment approach are the inability to incorporate the existing size composition information from the surveys and fishery. Therefore, the Collie-Sissenwine Analysis (CSA) model, which estimates the abundance and fishing and total mortality rates of recruit and post-recruit size classes, will be explored at the 2008 GARM in addition to an index-based assessment approach.

R. Gulf of Maine Haddock—Michael Palmer

1. Recent history of stock assessments
1986 - SAW 2 - No formal analysis of fishing mortality was attempted.

2001 - SAW 32 (NEFSC 2001a, b) - Availability of biological samples were insufficient to reliably estimate parameters needed to support a full analytic assessment of the stock. The stock determination was established using an exploitation rate computed as the three year centered average of the annual commercial landings divided by the annual Northeast Fisheries Science Center's (NEFSC) autumn survey stock biomass index. An update of the biological reference points was attempted using the ASPIC surplus production model (Prager 1994, Prager 1995), though the SARC Panel expressed concerns with the uncertainty of the model parameter estimates and recommended that previously established reference points be used (harvest control rule, ODRP 1998). The harvest control rule proxy exploitation rate index (F_{MSY}) and $B_{Threshold}$ (1/2 B_{MSY}) were estimated at 0.29 and 4.38 kg/tow, respectively. In 1999 the 3-year average exploitation rate index was 0.246 and the 3-year survey biomass index was 3.417 kg/tow. In 1999, the stock status was overfished, and overfishing was occurring. The Panel recommended that alternative estimates for deriving the exploitation index be investigated.

2002 – GARM (NEFSC 2002b) - Index based assessment based on the exploitation rate. Exploitation rate was computed as the annual commercial landings divided by the 3-year lagged average autumn survey stock biomass index. Biological reference points were determined using an index method (AIM, NEFSC 2002a) in response to concerns with the reference points calculated using the ASPIC surplus production model in the 2000 assessment. The proxy exploitation rate index (F_{MSY}) and $B_{Threshold}$ (1/2 B_{MSY}) were estimated at 0.23 and 11.09 kg/tow, respectively (NEFSC 2002a). The exploitation rate index in 2001 was 0.115 and 3-year survey biomass index was 10.31 kg/tow. In 2001, the stock status was overfished, but overfishing was not occurring.

2005 - GARM II (NEFSC 2005) - Index based assessment based on the exploitation rate. Exploitation rate was computed as the annual commercial landings divided by the 3-year average autumn survey stock biomass index. The exploitation rate index in 2004 was 0.18 and 3-year survey biomass index was 5.79 kg/tow. In 2004, the stock status was overfished, but overfishing was not occurring. The Panel recommended that future assessments include recreational catches and that application of age-structured models be investigated.

***2.* Strengths and Weaknesses**
The Gulf of Maine haddock stock is currently assessed using the index-based model, AIM. This approach is robust to the limited availability of accessory biological information (age, length, maturity, etc.) which has prohibited the use of alternative models in past assessments. The AIM model is sensitive to noise in the landings and survey time series, lack of discrimination in survey biomass between the recruited/pre-

recruit populations and requires a crude approximation of MSY to estimate a proxy for B_{MSY} (average of the commercial landings between 1959 and 1966, Clark et al. 1982).

3. Feasibility of Changing Assessment Models

Past GARM and SARC Panels have recommended investigating alternative models (VPA, statistical catch-at-age or length-based (e.g., SCALE)). There is limited biological data in many of the time series during the late 80s and early- to mid-90s which has prohibited applying these approaches in past assessments. Commercial (tables R.2, 3) and survey (table R.11) age and length data are limited between 1988 and 1995. Discard length samples are limited from 1989 to 1991 and 1997 to 2001 with no discard age samples available after 1999 (table R.6). Recreational length samples are limited prior to 2002 (table R.7). The primary goal for the GARM Models Meeting is to determine whether sufficient length and age samples exist to attempt alternative models on Gulf of Maine haddock, and if so, which model(s) is most appropriate.

References

Clark SH, Overholtz WJ, Hennemuth RC. 1982. Review and assessment of the Georges Bank and Gulf of Maine haddock fishery. J. Norhtw. Atl. Fish. Sci. 3:1-27.

Livingstone R, Jr. 1956. Conversion of total length to fork length for subdivision 5z haddock. In: Appendix to United States Research, 1956. p. 67.

Northeast Fisheries Science Center [NEFSC]. 2001a. Report of the 32nd Northeast Regional Stock Assessment Workshop (32nd SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Reference Document 01-05, Woods Hole, MA, 02543.

Northeast Fisheries Science Center [NEFSC]. 2001b. Assessment of 19 Northeast Groundfish Stocks Through 2000. NEFSC Reference Document 01-20, Woods Hole, MA, 02543.

Northeast Fisheries Science Center [NEFSC]. 2002a. Final Report of the Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. NEFSC Reference Document 02-04, Woods Hole, MA, 02543.

Northeast Fisheries Science Center [NEFSC]. 2002b. Assessment of 20 Northeast Groundfish Stocks through 2001. A report of the Groundfish Assement Review Meeting (GARM). NEFSC Reference Document 02-16, Woods Hole, MA, 02543.

Northeast Fisheries Science Center [NEFSC]. 2005. Assessment of 19 Northeast Groundfish Stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM). NEFSC Reference Document 05-13, Woods Hole, MA, 02543.

Overfishing Definition Review Panel [ODRP]. 1998. Evaluation of Existing Overfishing Definitions and Recommendations for New Overfishing Definitions to Comply with the Sustainable Fisheries Act. Overfishing Definition Review Panel, June 19, 1998 Final Report. 179 p.

Prager, MH. 1994. A suite of extensions to a non-equilibrium surplus production model. Fish. Bull. 92:374-389.

Prager, MH. 1995. User's manual for ASPIC: A stock production model incorporating covariates, program version 3.6X. Miami Laboratory Document MIA-92/93-55, National Marine Fisheries Service. 29 pp.

S. Atlantic Halibut—Laurel Col

1. History of Current Approach:

The status of the Gulf of Maine-Georges Bank Atlantic halibut stock is assessed based on research vessel survey indices and commercial catch. The halibut stock is classified as being overfished when the 5-year average of the NEFSC autumn survey swept-area biomass index falls below 2,700 mt, which is one-half of the B_{MSY} proxy of B_{MSY} = 5,400 mt (NEFSC 2002). Stock biomass has been relatively constant but has remained well below the overfished threshold since the 1960s. As a result, the stock has been in an overfished condition throughout the entire survey time series.

Overfishing on the Gulf of Maine-Georges Bank halibut stock occurs when fishing mortality exceeds the limit of F_{MSY}=0.06 (NEFSC 2002). Overfishing status cannot be determined because no estimates of fishing mortality are available. Nonetheless, exploitation rate indices (calculated as the annual catch divided by the 5-year average swept-area biomass) suggest that exploitation rates have probably been stable since the 1970s, and have declined to low levels since the early 1990s (Figure WP1.1 S4).

2. Model Selection:

Model selection for Atlantic halibut is severely limited by the available data. Due to the current low abundance of the stock, halibut are rarely encountered in either commercial landings or NEFSC surveys. The Gulf of Maine-Georges Bank Atlantic halibut stock collapsed roughly 100 years prior to the start of the NMFS fall survey, with landings peaking around 1860, and the available survey data does not capture the extent of the historical stock levels in contrast to current abundance. Although the time series of consistent landings starts just after the peak of the Atlantic halibut fishery, it does provide contrast and historical perspective. A stock reduction analysis, relying mainly on the commercial time series of data, may be a reasonable approach for assessing this data-poor stock (Figure WP 1.1 S5).